

# IFT6095 Robot Learning Homework 1

Lucas Maes

February 4, 2022

## 1 Behavior Cloning

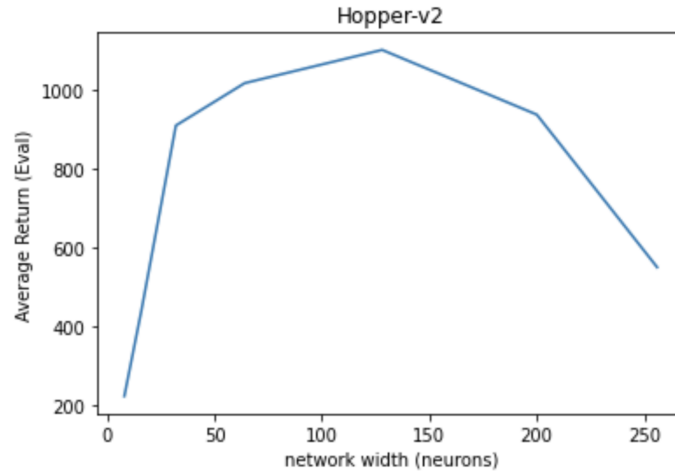
### 1.2

Environment	Expert mean	BC mean	% of Expert
Ant-v2	4713.65	3447.86 $\pm$ 1745	73%
Hopper-v2	3772.67	1047.92 $\pm$ 77.70	27,7%

**Table 1:** Expert and Deterministic Behavior cloning policies result on Ant-v2 and Hopper-v2. Network size: 64, Layers number: 2, Evaluation Batch size: 10000, Maximum episode length: 1000, Training steps per iteration: 1000

**note:** Maximum length of episode is 1000 and the evaluation batch size is 10000 that approximately 10 trajectories are collected.

### 1.3

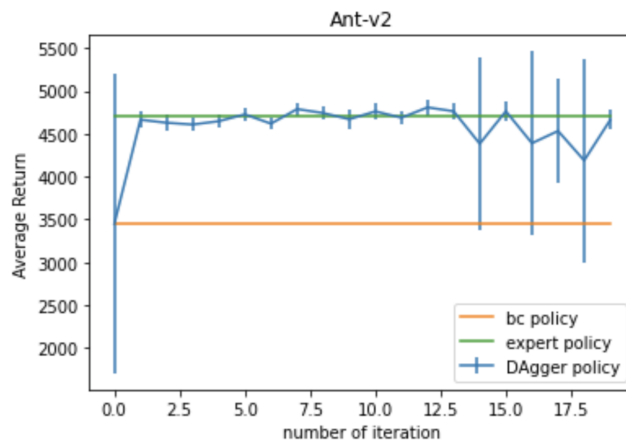


**Figure 1:** Average evaluation return as a function of network width on Hopper-v2.

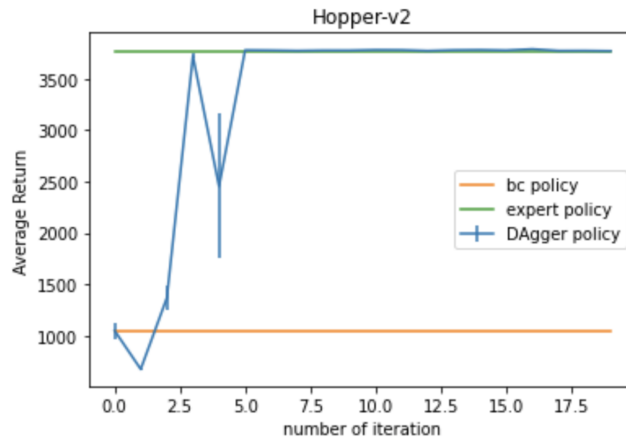
The network width was chosen because it is one of the factors that determine the capacity available on the policy network during the learning process. We can observe that the trend is quadratic and that a too large width of networks decreases the performance of the behavior cloning policy.

## 2 DAgger

### 2.2



**Figure 2:** BC policy, expert policy, and DAgger policy result on Ant-v2. Network size: 64, Number of iteration: 20, Layers number: 2, Evaluation Batch size: 10000, Maximum episode length: 1000, Training steps per iteration: 1000



**Figure 3:** BC policy, expert policy, and DAgger policy result on Hopper-v2. Network size: 64, Number of iteration: 20, Layers number: 2, Evaluation Batch size: 10000, Maximum episode length: 1000, Training steps per iteration: 1000